Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 7600–7604, August 1995
Biophysics

Significance of bound water to local chain conformations in protein crystals

CHARLES H. ROBERT* AND P. SHING HO

Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331

Communicated by John T. Edsall, Harvard University, Cambridge, MA, April 21, 1995 (received for review December 28, 1994)

ABSTRACT **We examine how the polypeptide chain in protein crystal structures exploits the multivalent hydrogen-bonding potential of bound water molecules. This shows that multiple interactions with a single water molecule tend to occur locally along the chain. A distinctive internal-coordinate representation of the local water-binding segments reveals several consensus conformations. The fractional water occupancy of each was found by comparison of the total number of conformations in the database regardless of the presence or absence of bound water. The water molecule appears particularly frequently in type II β-turn geometries and an N-terminal helix feature. This work constitutes a first step into assessing not only the generality but also the significance of specific water binding in globular proteins.**

Aqueous solvation has repeatedly revealed itself to be a fundamental factor in the stability and dynamics of macromolecules (e.g., refs. 1–4). As we continually improve our ability to observe the solvation of macromolecules (5–7), to describe its thermodynamics (2, 8), and to model solvent in atomic detail (e.g., ref. 9), we can expect to uncover more of water's influence in macromolecular behavior. It is natural to ask how this peculiar solvent interacts with proteins, but in this paper we turn the question on its head and ask how the polypeptide chain interacts with water. This exercise is not as formal as it may sound. Unlike small molecules, flexible polymers can adapt to the multivalent hydrogen-bonding capacity of water molecules by taking on a wide variety of conformations. How proteins optimize these interactions reflects on their native stability and on the folding process itself.

Studies of bound water in crystal structures have elucidated several themes. Water molecules tend to complete hydrogen-bonding patterns of individual chemical groups and of secondary structures (10, 11). In a series of homologous proteases, a conserved water molecule can be seen to replace an internal polar group to satisfy particular hydrogen-bonding requirements (12). Bound water can also be associated with various degrees of distortion of secondary structural elements, perhaps revealing a continuum of folding intermediates (13, 14), and combined crystallographic and calorimetric studies suggest that structurally integrated water molecules can stabilize the α-helix (15).

To place such observations into context, we have studied more generally how proteins in crystal structures incorporate water molecules into their folded states. We see that water complexation is typically a function of short sequence distances and, further, that the bridged polypeptide backbone tends to take on just a few preferred conformations. By examining how many such "water-competent" conformations are actually associated with water in the database, we discovered that those with the highest water occupancy are an N-terminal helix feature and type II β-turns. These conformations may exemplify structures whose hydrogen-bonding potentials are more difficult to satisfy internally.

METHODS

Crystal structures in the 1992 Brookhaven Protein Data Bank (16) were selected by a minimal number of criteria: within each group of related proteins identified by Walsh [Walsh, L. L., Brookhaven Protein Data Bank annotation, April 1992], only the highest-resolution entry with a nonzero number of waters and resolution better than 2.5 Å was selected. In the resulting set of protein structures (Table 1), we investigated only monomeric proteins, first-chain entries of multimeric proteins, and the enzyme part of enzyme–ligand complexes. Entries from different organisms but with only a few amino acid differences were manually edited from the final set of proteins.

We constructed a spherical shell with a radius of 3.2 Å about each water oxygen and gathered the coordinates of all fully occupied main-chain and side-chain C=O and N—H hydrogen-bonding groups within the shell. Each candidate hydrogen-bonding group was given characteristic "target" binding locations for the water oxygen atom. These loci were assigned by the geometries of Thanki *et al.* (7) at an averaged $O_W \cdots X$ distance of 2.9 Å (where w is water). We reduced the set of candidate hydrogen-bonding partners by requiring in addition that one of these target loci falls within the shell surrounding the water oxygen. Of course if the group were canonically hydrogen-bonded to the water molecule, its target locus would coincide with the water oxygen at the center of the shell. These simple dual criteria avoid the need for a hard cutoff of the hydrogen-bonding angle while automatically excluding extremely nonlinear hydrogen-bonding geometries (for example, with typical van der Waals radii the most extreme NH···O angle is ≈90°). The chain segments defined by the first and last interaction with a given water molecule were transformed to the coordinate system of the first candidate by our program and written to files for further analysis. These will be referred to as fragments.

For each fragment, we tabulated the amino acid frequency at each position and calculated ϕ, ψ, ω angles and through-space distance between first and last residues. We used the Connolly accessible-surface program (17) with extended van der Waals radii to estimate solvent exposure of the 20 amino acids and the chain fragments in their protein contexts (excluding water). Finally, we systematically searched our data base for all occurrences of a given conformation to estimate fractional water occupancy. Temperature factors of the water molecules found in the conformational searches were at or below the average water values in the protein in >80% of the cases. Searches performed by using crystal structures with only 2.0-Å resolution or better resulted in essentially unchanged occupancies. All programs were written with FORTRAN 77 on a Silicon Graphics Iris workstation or with MATHEMATICA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Present address: Laboratoire de Biologie Structurale, 91198 Gif-sur-Yvette, Cedex, France.

Table 1. Protein Data Bank structures grouped by occurrences of 3- to 5-residue fragments

Total fragments, no.	Protein Data Bank identifier
0	*2abx*, *8adh*, *1cd4*, *1fxi*, *3gap*, **1gma**, *2gn5*, **2hmz**, 1hoe, **4ins**, *2liv*, 2mlt, **2ovo**, *1phh*, **4pti**, 1r69, **1utg**, **1xyl**
1	*1cdt*, *2cro*, *1dtx*, **3ebx**, 2fd2, **3fgf**, **1fkf**, 5hvp, *3icb*, *7icd*, *1msb*, *1mvp*, 1omd, **1pal**, **9pap**, 1rnh, **3rn3**, **2sga**, 1sn3, **451c**, **256b**
2	**2alp**, **2ccy**, **2csc**, **1ctf**, **3c2c**, 4i1b, 1ifb, 7pcy, **5p21**, **2rhe**, **4rxn**, **1snc**, *2stv*, 2tsc, 1ubq
3	**1aap**, **1alc**, 3blm, **1bp2**, 2ca2, 2cdv, **4cpv**, **1gcr**, 2lzt, **1paz**, **4ptp**, 3rnt, 3rp2, 3sgb, **1ycc**
4	2act, *8atc*, 2aza, *8cat*, **5cpa**, 1hne, **3lzm**, *4mdh*, *2pfk*, 1rbp, 2tec, **1thb**
5	1bbp, 3bcl, *4cms*, 1cox, **1eco**, **3gct**, 6ldh, 7wga
6	**3b5c**, **2cpp**, **3dfr**, 2gbp, 1gp1, **1mbd**, **5rub**, 2tim 5tnc, 1ton, *2ts1*
7	**2er7**, *1fcb*, **2ltn**, **2prk**, **2trx**
8	**8abp**, **2cyp**, **1psg**, **5tmn**
9	**6xia**
≥10	*5acn*, 3cla, **1cse**, 1gox, **3grs**

Boldface, normal, and italic types are used to indicate structure resolution ≤1.75 Å, 1.75 < resolution ≤ 2.0 Å, and 2.0 < resolution ≤ 2.5 Å, respectively.

(Wolfram Research, Champaign, IL) on a Macintosh IIci computer.

RESULTS AND DISCUSSION

Distribution of Sequence Distances. Fig. 1 shows how the sequence distance between the first and last water interaction is distributed in the database. The distribution is clearly peaked at short sequence distances. The effects of side-chain donors and acceptors appear largely limited to within one residue of the first interaction, suggesting that it is backbone conformational characteristics that govern the chain's potential for forming multiple interactions. We restrict ourselves to main-chain donors and acceptors for the remainder of this paper. The rapid dropoff of the distribution with sequence distance highlights the local nature of water interactions in proteins; indeed, 40% of the bridging interactions are within 5 residues. Stickle *et al.* (18) found a similar but more marked preference for local intramolecular hydrogen bonding. We also examined candidate pairs that qualified for a water-bridging interaction but that were gathered without presupposing a water molecule (data not shown). This distribution showed fewer long-range pairs, indicating that bound water links distant elements in the sequence in addition to local residues.

Classes of Fragments. We focus on the short (3–5 residue) water-binding fragments. The number found in each protein can be seen in Table 1. The median is two-to-three fragments per protein; 97 of the 115 proteins have at least one fragment. Extracted fragments are grouped into four categories or classes, describing the four possible water-bridged interactions between the first and last residues: Classes: a, C═O···w···O═C; b, C═O···w···H—N; c, N—H···w···H—N; d, N—H···w···O═C. Table 2 shows the number of each class extracted from the database.

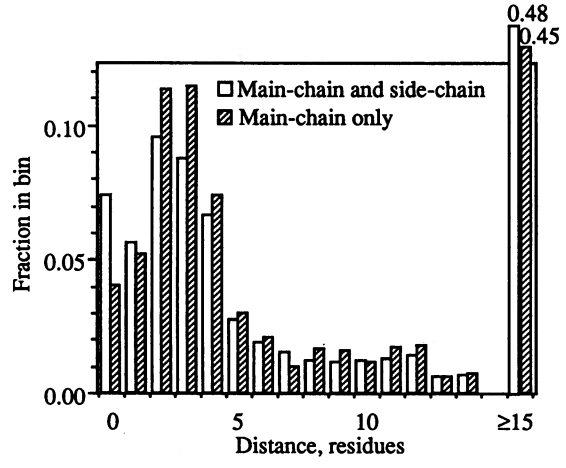

FIG. 1. Distribution of sequence distance between first and last water interaction along the polypeptide chain for main-chain candidates alone and for main-chain and side-chain candidates, normalized to the total number of occurrences of each type.

Surface Accessibilities. We expected that a high proportion of the fragments extracted in this study would be situated close to the protein or subunit surface. One reason for this is the relative rarity of water bound in the protein interior. A second is that the chain will tend to bend back on itself to form the second hydrogen bond with the water molecule, resulting in a turn. Such turns are known to occur at hydrophobicity minima and to occur with high frequency at the protein surface (19). Fig. 2 shows that most of the fragments demonstrate significant surface exposure. In fact, ≈90% of the fragments fall into a single distribution with ≈6 Å^2 atom on average accessible to our probe.

Amino Acid Compositions. The fragments thus appear to be primarily located near the protein surface. An enrichment of certain amino acids, most obviously for hydrophilic residues, is expected for this reason alone. We wish to separate out this contribution to the fragment compositions. By using the same methods used for fragment exposures, we estimated surface exposures of the 20 amino acids in our database. The resulting distributions showed well-separated buried and surface populations that could be delimited by an average probe exposure

Table 2. Classes of water-binding fragments with amino acid enrichment

Fragment class (*N*)	Type	Residue 1	2	3	4	5
3a (73)	O···w···O	G_{12},C_6	W_5	G_{12}		
3b (57)	O···w···N	G_{13}	D_{12},I_6	M_4,T_8		
3c (5)	N···w···O					
3d (3)	N···w···N					
4a (83)	O···w···O	—	D_{12}	G_{15}	V_{12},F_9	
4b (53)	O···w···N	—	C_4,G_9	G_9	—	
4c (1)	N···w···O					
4d (7)	N···w···N					
5a (27)	O···w···O	L_6	E_5	—	G_{12}	G_6
5b (65)	O···w···N	D_{10}	M_4	L_9	L_9	A_{10}
5c (0)	N···w···O					
5d (12)	N···w···N					

Significances (≥95%) in major fragment classes were calculated by discrete Poisson distribution; $Np \leq 6$. Subscripts give number of occurrences in this class. —, Positions at which no significant enrichment was seen.

of 1.2 Å^2 per atom, simultaneously distinguishing surface from buried fragment populations in Fig. 2. The resulting propensities allowed us to estimate (at a given confidence) the maximum number of occurrences of each amino acid we would expect in each set of extracted fragments. Any amino acid with greater than this number of occurrences is likely to be present for some reason beyond its surface propensity (Table 2).

Our fragments were selected through their main-chain interactions, although there is no reason that side-chain interactions could not help enrich polar amino acids. But indeed the amino acid identities appear more likely to be felt through their backbone steric properties. For instance, we see an overall preference for glycine in many of the fragments in Table 2. This residue's nonpenalizing steric map provides a maximal possibility for the chain to fold back on itself to achieve a second water interaction. In the last fragment class (class 5b), the enriched amino acids are consistent with α-helical structures, as will be discussed below.

ϕ,ψ Angle Distributions. Next we examined individual ϕ,ψ plots for each residue position in the extracted fragments. Clustering of ϕ,ψ values demonstrated that the extracted chains reside in well-defined conformations. For example, the 3-residue fragments are shown in Fig. 3. For such small fragments, it is understandable that there are only one or two common ways to achieve a bridging interaction. Yet for 4- and 5-residue fragments, the appearance of clustering was surprising; we expected that such fragments would be found to complex the water molecule in many different ways.

Stacked ϕ,ψ Plots. A problem with viewing any collection of polypeptide chain fragments in terms of individual ϕ,ψ plots is that the correlation among residues for each individual chain is lost. Alternatively, examining a set of fragments in a common three-dimensional space allows one to see each structure clearly, but this approach suffers from the arbitrary choice of a coordinate system for each fragment. To circumvent these problems, we introduced the stacked ϕ,ψ planes shown in Fig. 4. In these three-dimensional plots, each residue is plotted in the ϕ,ψ plane corresponding to its position in the fragment. A residue is connected by lines to its neighboring residues in adjacent planes. The structure of each chain fragment is then seen by following along the line from one plane to the next. If the sample has a consensus conformation, a ropy data structure becomes obvious.

Conformational Motifs. We made use of both conventional and stacked ϕ,ψ plots to delimit the major conformations seen in each fragment class. Ranges for the internal coordinates and end-to-end distances of the major structural motifs are presented in Table 3.

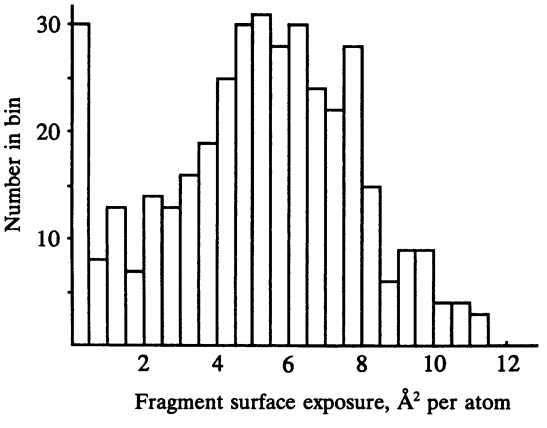

FIG. 2. Approximate values for the accessible surface of extracted fragments, calculated by using the Connolly algorithm (14) with extended van der Waals radii (2.1, 1.9, 2.05, 2.2, and 2.05 Å for C, O, N, S, and heteroatoms, respectively) and a probe radius of 2.5 Å.

The 3-residue fragment classes are markedly distinct, each characterized by a single cluster of ϕ,ψ values (Fig. 3). The conformation of fragments in class 3a (O···w···O) is consistent with the first residues of an opened-up type I β-turn, a motif echoed in the larger fragments as well. However, the conformation of fragments in class 3b (O···w···N) resembles a γ-turn, which is defined by a C=O$_{(1)}$···H—N$_{(3)}$ hydrogen bond. Sundaralingam and Sekharudu (20) also report a prevalence of γ-turns associated with water molecules in crystal structures. Milner-White (21) surveyed γ-turns in crystal structures and classified them according to Ramachandran plots for the second residue. According to his mapping of a hydrogen-bond strength measure onto the ϕ,ψ plane, those fragments selected in class 3b would be among the more stable γ-turns.

The two classes of 4-residue fragments (classes 4a and 4b) are closely related, despite the different manners in which they are bridged by the water molecule. Both contain two major populations of fragment conformations; a third is visible in class 4b. Across the classes, the major (classes 4a1 and 4b1) and the minor (classes 4a2 and 4b2) conformations are similar as well. The major conformation is close to the common type I β-turn, which is often found to be stabilized by an intramolecular C=O$_{(1)}$···H—N$_{(4)}$ hydrogen bond across the turn (22). In our fragments, the carbonyl of the first residue points somewhat out of the plane of the turn, where it is more accessible for the bridging interaction with the water molecule, and is more aligned in the direction of a helical hydrogen bond. Model peptides demonstrating a similar conformation have been used to support the idea of β-turns as intermediates in the formation of helical structures (23). The fragments in conformation 4b1 are somewhat farther from the helix region than those in conformation 4a1. They are like type I turns that have been opened still further. The last 2 residues begin a turn in the opposite direction, placing a slight jog in the chain path. Despite their distance from the canonical α-helical conformation, we found that fragments with this conformation were frequently assigned to helices. Sekharudu and Sundaralingam (14) suggest that these and nearby conformations in the ϕ,ψ neighborhood may manifest encroachment by water molecules into the helix.

The less-populated conformations (4a2 and 4b2) are closer to type II β-turns. The carbonyl of the second residue is antiparallel to that of the first, which essentially precludes any role as an intermediate in the formation of an α-helical structure. They are smoother and more planar than the type I turns discussed above. Fragments with these conformations were often found to be unassigned with respect to secondary structural classes. As will be seen below, the water molecule is

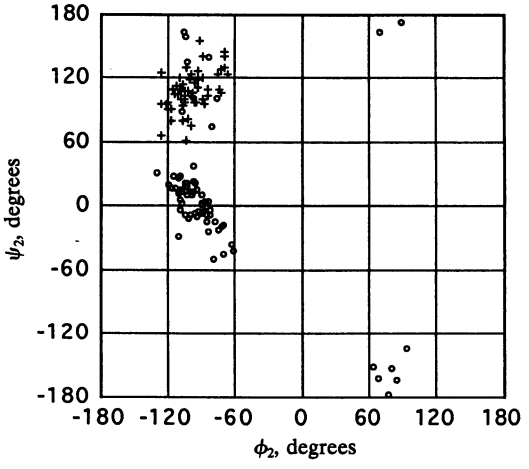

FIG. 3. ϕ,ψ plots for the interior residue of 3-residue fragments: classes 3a (○) and 3b (+).

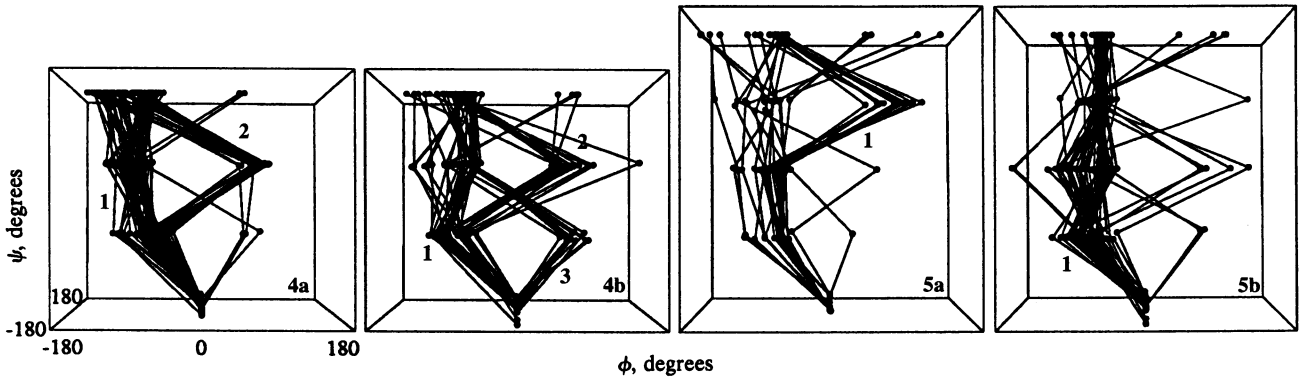

FIG. 4. Stacked ϕ,ψ plots for 4- and 5-amino acid fragments revealing consensus conformations (numbered) of water-binding fragments. For the first residue, ϕ is assigned to zero; for the last residue, ψ is assigned to zero.

also more closely tied to the occurrence of these conformations in the database.

Each class of 5-residue fragments appears to have only a single conformation overall. In class 5a, the initial residues are approximately helical. Indeed, when we examined secondary structure assignments, we found that all the fragments with conformation 5a started within 2 residues of the C-terminal end of a helix. The predominance of glycine in the last two positions (Table 2) is consistent with helix termination. These water-bridged conformations thus make up a C-terminal helix-capping motif. In conformation 5b, on the other hand, 19 of 22 assigned fragments fell within 3 residues of the N terminus of a helix, and the amide partner was within an assigned helical region in 95% of the cases. The deviations of the residues from helicity are not as marked as in conformation 5a, and in a few cases, these fragments formed kinks in a longer helix.

Fractional Water Occupancies. When a stretch of the polypeptide chain is found to be in one of the conformations we have just seen, how often is there a water molecule associated with it? To answer this question, we searched through the database once again, this time to find all runs of n contiguous amino acids whose structural parameters fell within the allowed regions for each conformation. For consistency we searched the fragment sets with the same criteria. The number of structures found in these searches and calculated occupancies are given in Table 3. As one would expect from the distinctive clustering seen in the stacked ϕ,ψ plots, the number of occurrences found in the extracted fragments is significantly higher than that expected statistically from their occurrence in the database as a whole.

While the bound water is significant to all of the major fragment conformations seen in our study, it is not immediately obvious why it is far more important to certain structures than to others. However, we suggest that these conformations must be considered in the context of the entire protein. For example, conformations 4a1 and 4b1 are close to the commonly occurring type I turns. These features can be associated with distorted or water-inserted helices (13), and consequently, they have a higher chance of being couched in that extensive backbone hydrogen-bonding network. But the type II conformations (4a2 and 4b2) are more likely to lie in a less-structured environment—thus, the bound water may contribute proportionately more to their stabilization.

Similarly, in the 5-residue fragments, both of the major conformations can be reasonably called water-associated helix-capping features, yet the water is far more common at the N terminus (class 5b) than at the C terminus (class 5a). This finding is remarkably similar to the preference for side-chain capping at the N terminus of helices (24) and is also consistent with the observation that impeding N-terminal solvation results in a measurable decrease of helix stability (15). One possibility to explain these observations would be the more precise geometry required by the amide donor (7, 10) at the start of the helix than by the corresponding carbonyl at the far

Table 3. Conserved conformations of polypeptide-chain fragments and fractional water occupancies

		Range of residue ϕ/ψ conformational angles, degrees										$C_{\alpha 1}$–$C_{\alpha n}$ distance, Å		Number found		%
		1		2		3		4		5						
Conformation		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Self	Global	occupancy
3a	ϕ			−120	−72	−96	−48					5.0	7.0	18	409	4.4
	ψ	−36	0	−24	24											
3b				−132	−84	−108	−48					5.5	7.0	22	775	2.8
		108	156	84	132											
4a1				−84	−48	−120	−72	−156	−60			4.5	7.5	23	526	4.4
		−48	−12	−48	0	−36	24									
4a2				−84	−36	72	96	−156	−60			4.5	7.5	10	47	21
		108	168	120	156	−20	30									
4b1				−132	−60	−84	−36	−108	−48			6.5	8.5	8	98	8.2
		−60	0	−20	36	−60	0									
4b2				−132	−60	48	108	−108	−48			6.5	8.0	5	25	20
		84	156	120	180	0	50									
5a				−72	−48	−96	−48	48	132	−108	−48	4.5	6.5	5	83	6.0
		−60	−24	−60	0	−48	24	−12	36							
5b				−120	−36	−132	−36	−96	−36	−84	−36	7.5	9.0	23	111	21
		−60	0	−60	−12	−72	12	−72	−24							

Self, number of water-associated n-residue fragments occurring with given conformation. Confidences ≥95%, evaluated as in Table 2. Global, number of all n-residue fragments in database occurring with given conformation.

end. Backbone polar groups, which are the preferred hydrogen-bonding partners (25, 26), would not be as free to meet the geometry conditions at the N terminus, so the proportion of these hydrogen-bonding potentials satisfied by flexible side chains and/or water molecules would increase.

Conclusions. Of the widespread occurrence in protein crystal structures of hydrogen-bond donors and acceptors bridged by water molecules, nearly half occur within 5 residues of one another. This preference for local binding appears to be a general property of the polypeptide backbone. The extent to which specifically bound water stabilizes the fragment conformations found in this study cannot yet be estimated, since neither the strength of the interaction nor the stability of the chain structure is known; however, even a small advantage would contribute at the level of local structure formation. We have seen that the bridging water molecule is especially implicated in type II β-turn conformations and N-terminal helix structures. The hydrogen-bonding potentials of these conformations may be more difficult to satisfy in globular proteins; if so, they could be sensitive in the balance of folding and unfolding tendencies.

We thank Robert Martin for help in the calculation of amino acid surface accessibilities for our database, Dr. Gary P. Schroth for discussion, and referee 2 for suggesting the analysis of candidate sequence distributions. C.H.R. also thanks Kensal van Holde for his continued support throughout this work. This work was funded by Grant MCB9304467 from the National Science Foundation (P.S.H.), Grant N00014-91-J-1542 from the Office of Naval Research (P.S.H.) and by the John Simon Guggenheim Memorial Foundation (C.H.R.).

1. Levitt, M. & Park, B. H. (1993) *Structure* **1,** 223–226.
2. Privalov, P. L. & Makhatadze, G. I. (1993) *J. Mol. Biol.* **232,** 660–679.
3. Steinbach, P. J. & Brooks, B. R. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 9135–9139.
4. Kagawa, T., Stoddard, D., Zhou, G. & Ho, P. S. (1989) *Biochemistry* **28,** 6642–6651.
5. Otting, G., Liepinsh, E. & Wütrich, K. (1991) *Science* **254,** 974–980.
6. Teeter, M. M. (1984) in *Protein Folding: Deciphering the Second Half of the Genetic Code*, eds. King, J. & Gierasch, L. (Am. Assoc. Adv. Science, Washington, DC), pp. 44–54.
7. Thanki, N., Thornton, J. M. & Goodfellow, J. M. (1988) *J. Mol. Biol.* **202,** 637–657.
8. Ben-Naim, A. (1987) *Solvation Thermodynamics* (Plenum, New York).
9. Brooks, C. L., III, Karplus, M. & Pettitt, B. (1988) *Adv. Chem. Phys.* **71,** 137–174.
10. Thanki, N., Umrania, Y., Thornton, J. M. & Goodfellow, J. (1991) *J. Mol. Biol.* **221,** 669–691.
11. Roe, S. M. & Teeter, M. M. (1993) *J. Mol. Biol.* **229,** 419–427.
12. Sreenivasan, U. & Axelsen, P. H. (1992) *Biochemistry* **31,** 12785–12791.
13. Sundaralingam, M. & Sekharudu, C. Y. (1989) *Science* **244,** 1333–1337.
14. Sekharudu, C. Y. & Sundaralingam, M. (1993) in *Water and Biological Macromolecules*, ed. Westhof, E. (CRC, Boca Raton, FL).
15. Harpaz, Y., Elmasry, N., Fersht, A. R. & Henrick, K. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 311–315.
16. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112,** 533–542.
17. Connolly, M. (1981) Ph.D. thesis (Univ. of California, San Francisco).
18. Stickle, D. F., Presta, L. G., Dill, K. A. & Rose, G. D. (1992) *J. Mol. Biol.* **226,** 1143–1159.
19. Rose, G. D. (1978) *Nature (London)* **272,** 586–590.
20. Sundaralingam, M. & Sekharudu, C. (1994) *Biophys. J.* **66,** 346 (abstr.).
21. Milner-White, E. J. (1990) *J. Mol. Biol.* **216,** 385–397.
22. Chou, P. Y. & Fasman, G. D. (1977) *J. Mol. Biol.* **115,** 135–175.
23. Perczel, A., Foxman, B. M. & Fasman, G. D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8210–8214.
24. Harper, E. T. & Rose, G. D. (1993) *Biochemistry* **32,** 7605–7609.
25. Baker, E. N. & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol.* **44,** 97–179.
26. Aurora, R., Srinivasen, R. & Rose, G. D. (1994) *Science* **264,** 1126–1130.